Evaluating *Z* Scores to Quantify Levator Hiatal Distensibility by 3-Dimensional Ultrasonography in Nulliparas and Women With Pelvic Organ Prolapse

Lieming Wen, MD, Minghui Liu, PhD, Baihua Zhao, MD, Zhenzhen Qing, MD

Objectives—To use *Z* scores to quantify hiatal distensibility and to test the performance of *Z* scores for levator hiatal areas in predicting substantial pelvic organ prolapse (POP).

Methods—We undertook a retrospective study of the data from 145 nulliparas and 166 patients with POP who had a clinical POP examination with 3-dimensional translabial ultrasonography. *Z* scores were used to normalize levator hiatal areas of nulliparas. The *Z* score model for the hiatal area was built by the formula *Z* score = (measured value – predicted mean value)/predicted standard deviation and was used to evaluate hiatal ballooning in women with POP.

Results—Valid data were gathered from 134 nulliparas and 159 patients with POP. POP stage 1 was found in 46 women, stage 2 in 62, stage 3 in 43, and stage 4 in 8. We built the *Z* score model as follows: *Z*-Av = (measured value – 17.15)/3.11, where Av represented the minimal levator hiatal area on the maximum Valsalva maneuver. The levator hiatal area was strongly related to the POP stage ($P < .001$). On a receiver operating characteristic curve analysis, the cutoff of *Z*-Av was 1 for POP stage 2 or higher (sensitivity, 77%; specificity, 60%) and substantial POP on ultrasonography (sensitivity, 84%; specificity, 75%).

Conclusions—Hiatal distensibility can be exactly evaluated by *Z*-Av. A *Z*-Av value of less than 1.0 was defined as a "normal hiatal expansion," 1 to 3 as "mild ballooning," 3 to 5 as "moderate ballooning," 5 to 7 as "marked ballooning," and 7 or greater as "severe ballooning."

Key Words—hiatal ballooning; levator hiatus; pelvic floor; pelvic organ prolapse; 3-dimensional ultrasonography; *Z* score

Received October 20, 2017, from the Department of Ultrasound Diagnosis, Second Xiangya Hospital, Central South University, Changsha, China. Manuscript accepted for publication December 19, 2017.

This study was supported by the Project of New Clinic Techniques of Central South University.

Address correspondence to Minghui Liu, PhD, Department of Ultrasound Diagnosis, Second Xiangya Hospital of Central South University, 139 Renmin Rd M, 410011 Changsha, Hunan, China.

E-mail: liuminghui@csu.edu.cn, liuminghui1203@outlook.com

Abbreviations
Ar, levator hiatal area at rest; Av, levator hiatal area on the maximum Valsalva maneuver; ICS, International Continence Society; POP, pelvic organ prolapse; POP-Q, Pelvic Organ Prolapse Quantification; ROC, receiver operating characteristic; US, ultrasonography

doi:10.1002/jum.14590

The levator hiatus is the largest portal in the abdominal envelope. Levator hiatal biometric characteristics and function are strongly related to pelvic organ prolapse (POP).[1–4] Hiatal ballooning (ie, extensive hiatal expansion), had an evident association with POP and POP recurrence.[1–5] Translabial ultrasonography (US) was reported as a valid, practicable means to measure the levator hiatal dimensions and has seen widespread use for more than 20 years.[6] Dietz et al[4,6] proposed that extensive hiatal distensibility was related to POP stage 2 or higher: ie, hiatal ballooning was determined by measuring the levator hiatal area on the Valsalva maneuver; a measurement of 25 to 30 cm^2 was defined as "mild ballooning," 30 to 35 cm^2 as "moderate ballooning," 35 to 40 cm^2 as "marked

© 2018 by the American Institute of Ultrasound in Medicine

ballooning," and 40 cm^2 or greater as "severe ballooning." However, previous studies demonstrated that the levator hiatal area varied greatly in the population and had considerable ethnic differences. Albrich et al[7] produced a cutoff value of 27.53 cm^2 for substantial hiatal ballooning in women with lower urinary tract symptoms. A study of 52 nulligravid white women showed that the levator hiatal area ranged from 6.34 to 18.06 cm^2 at rest and from 6.67 to 35.01 cm^2 on the Valsalva maneuver.[3] Shek KL et al[8] found that the hiatal area had statistically significant differences between Ugandan and white women (mean levator hiatal area at rest $\pm$ SD, 15.66 $\pm$ 3.67 versus 11.39 $\pm$ 2.46 cm^2).

The *Z* score, which is widely used in biometry, has been described as the measured value that deviates from the average level as a multiple of the standard deviations. It can be used to precisely quantify the difference between the individual value and the general level.[9–11] In statistics, the *Z* score (ie, standard score) is the signed number of standard deviations by which the value of an observation or data point is above or below the mean value of what is being observed or measured. Observed values above the mean have positive standard scores, whereas values below the mean have negative standard scores. If a *Z* score is 0, it indicates that the score is identical to the mean score. This *Z* score conversion process is called standardizing or normalizing, which does not change the data distribution.

Using *Z* scores to assess the levator hiatal area allows for the exact quantification of the individual differences in the hiatal area and the degree of hiatal ballooning. In a previous study, we successfully used *Z* scores to analyze the hiatal dimensions of 102 Chinese nulliparas and explored the relationship between the hiatal dimensions and POP. The *Z* score models for the hiatal area and anteroposterior diameter were built to calculate the hiatal expansion degrees of 64 women with POP.[12] In this study, we further tested the performance of the *Z* scores for the hiatal area in predicting substantial POP clinically and on US.

Materials and Methods

We undertook a retrospective observational study of the data from 145 Chinese nulliparas who were recruited in a previous study for normal levator hiatal biometric measurements and 166 patients with POP who were seen for lower urinary tract symptoms or POP between January 2016 and May 2017 at the Second Xiangya Hospital. The inclusion criteria were as follows: (1) all of the nulliparas had no prolapse symptoms and signs; (2) all women had no history of pelvic or pelvic floor surgery or physiotherapeutic interventions for a pelvic floor disorder; (3) all patients with POP had a record of a standard clinic interview with POP clinical examinations performed according to the Pelvic Organ Prolapse Quantification (POP-Q) system of the International Continence Society (ICS)[13] by a specially trained gynecologist; and (4) all women had 3-dimensional translabial US volume data.

Translabial US was performed on patients in the supine position after bladder emptying with a Voluson 730 Expert or E8 system (GE Healthcare, Milwaukee, WI) equipped with a 4–8-MHz curved array volume transducer that was covered with a condom and placed on the perineum in a sagittal direction with an 85° acquisition angle. The operator took care to use minimal pressure to ensure maximal POP. Volume acquisition was performed at rest and on the maximum Valsalva maneuver and pelvic floor muscle contraction.[3,6] At least 3 Valsalva maneuvers and pelvic floor muscle contractions were completed per patient. Postprocessing of the US volume data was performed several months later with 4D View version 10.3 software (GE Healthcare), and the researchers were blinded to all of the clinical data. The data obtained in the context of this study were approved by the Human Research Ethics Committee of the Second Xiangya Hospital. All data were obtained by informed patient agreement.

The most effective (in terms of producing organ descent) Valsalva maneuver was chosen to perform the measurement. Ultrasonographic measurements of organ descent (bladder, uterus, rectal ampulla, or enterocele) were performed against a horizontal line placed through the inferior symphyseal margin, giving the maximal caudal position on the Valsalva maneuver. The levator hiatal area was measured on the axial plane with minimal dimensions, which was determined by the minimal hiatal anteroposterior diameter: ie, the shortest distance between the hyperechoic inferoposterior aspect of the pubic symphysis and the hyperechoic anterior border of the pubovisceral muscle that is just posterior to the anorectal muscularis in the midsagittal plane (Figure 1).[3,6] The intra-class correlation coefficient of a test-retest series of 20 patients was 0.83 (95% confidence interval, 0.77–0.95).

Prolapse symptoms were defined as a vaginal lump/bulge or a dragging sensation. Substantial POP on the clinical examination was defined as POP-Q stage 2 or higher.[4] Substantial POP on translabial US was defined as a cystocele to at least 10 mm below the symphysis, uterine descent to 15 mm above the symphysis or lower, rectal ampulla/enterocele descent to at least 15 mm below the symphysis, or a combination thereof.[14]

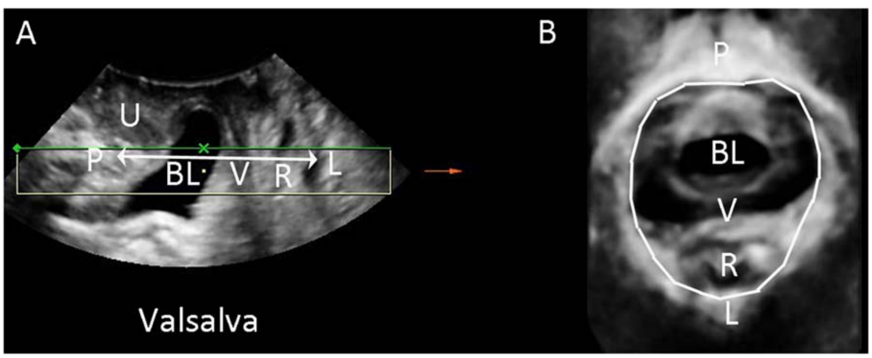
Descriptive statistics for demographic, US findings, and clinical data were obtained with Microsoft Excel (Microsoft Corporation, Redmond, WA). The statistical analysis was undertaken with SPSS version 17.0 software for Windows (IBM Corporation, Armonk, NY). The measurements were transformed into Z scores according to the Z score model for the levator hiatal area on the Valsalva maneuver, which was built on the basis of the nulliparous population by using the method documented in a previous study.[10] The procedures for building the Z score model are described below. First, data were tested for a normal distribution by a Kolmogorov-Smirnov analysis. Second, a Pearson correlation coefficient analysis was performed to correlate the measurements at rest and on the Valsalva maneuver. The correlation was considered statistically significant at $r > 0.7$ with $P < .01$ (2 tailed). Third, we ran a logistic regression analysis to build the linear regression equations and calculated the predicted value and standard deviation. Fourth, we built a Z score model using the following formula: Z score = (measured value − predicted mean value)/predicted standard deviation. Z score values were tested for a normal distribution by a Kolmogorov-Smirnov analysis. Finally, we determined the 95% reference range of the Z scores. A receiver operating characteristic (ROC) curve analysis was used to evaluate the association between Z scores and substantial POP clinically and on US in women with POP. $P < .05$ was considered statistically significant.

Results

Seven women with POP and 11 nulliparas were removed because of incomplete volume data. Valid data were gathered from 134 nulliparas, with a mean age of 32 (range, 20–41) years and a mean body mass index of 21.20 (range, 17.21–31.65) kg/m^2, and 174 women with POP, with a mean age of 48 (range, 35–82) years and a mean body mass index of 22.53 (range, 18.35–26.52) kg/m^2.

All values for the 134 nulliparas were normally distributed. The mean levator hiatal areas ± SD at rest and on the Valsalva maneuver were 13.18 ± 2.40 (range, 8.38–18.66) and 17.14 ± 3.45 (range, 10.23–26.14) cm^2, respectively. The Pearson correlation coefficient analysis revealed an excellent correlation between the levator hiatal area at rest and on the Valsalva maneuver ($r = 0.892$; $P < .001$). The regression equation for levator hiatal area on the Valsalva maneuver was established as $Av = 0.096 + 1.294 \times Ar$, where Av and Ar represented the levator hiatal area on the Valsalva maneuver and at rest, respectively. A scatterplot diagram presents this relationship in Figure 2. We calculated the predicted value for Av according to the regression equation and

Figure 1. Three-dimensional translabial US of a woman with substantial cystocele showing measurement of the hiatal area. **A**, On the maximal Valsalva maneuver, minimal axial view determined on the midsagittal plane by the minimal hiatal anteroposterior diameter: ie, the shortest distance between the hyperechoic inferoposterior aspect of the pubic symphysis and the hyperechoic anterior border of the pubovisceral muscle (double-sided arrow). **B**, Measurement of the hiatal area (white outline) on the minimal axial plane. BL indicates bladder; L, levator ani; P, symphysis pubis; R, rectum; U, urethra; and V, vagina.



built the Z-Av model as Z-Av = (measured value – 17.15)/3.11 using the Z score formula. A diagram for the Z score transformation is shown in Figure 2.

The Av values were transformed into Z scores, which were normally distributed, as confirmed by Kolmogorov-Smirnov analysis. The mean of Z-Av was 0, with an SD of 1.12. The 95% reference range of Z-Av was between −2.0 and +2.0, which was calculated by the following formula: 95% reference range = mean ± 1.96 × SD.

Of the 159 women with POP, 46 (28.9%) had POP stage 1; 62 (38.9%) had POP stage 2; 43(27.0%) had POP stage 3; and 8 (5.0%) had POP stage 4. One hundred fourteen (71.6%) patients had a diagnosis of substantial POP on translabial US, which was a cystocele in 58.4% (n = 93), uterine prolapse in 34.0% (n = 54), a rectocele in 15.1% (n = 24), and an enterocele in 3% (n = 5). Forty-four (96%) patients with POP stage 1 and 11 (18%) with POP stage 2 had no POP symptoms. All 51 (100%) patients with POP stage 3 or higher had POP symptoms, including a dragging sensation, a vaginal lump, and lower urinary tract symptoms (ie, leakage, urgency, frequency, nocturia, or voiding dysfunction).

The measured Av values were transformed into Z scores by using the Z score model built above. The measurements and Z scores were analyzed and showed similar outcomes. The values were significantly related to the POP stage ($r = 0.694$; $P < .01$), and the mean values increased as the POP-Q stage increased (Figure 3). An analysis of variance showed that there were

Figure 3. Error bar chart showing the mean Z scores for the hiatal area on the Valsalva maneuver and 95% confidence intervals according to the ICS POP stages.

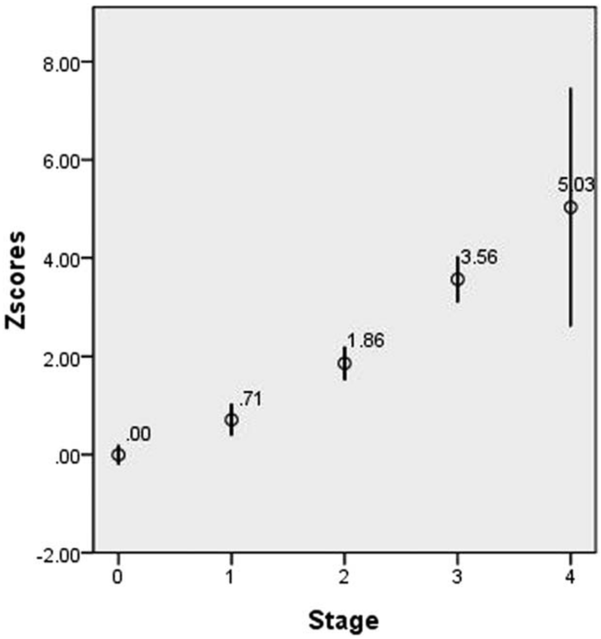
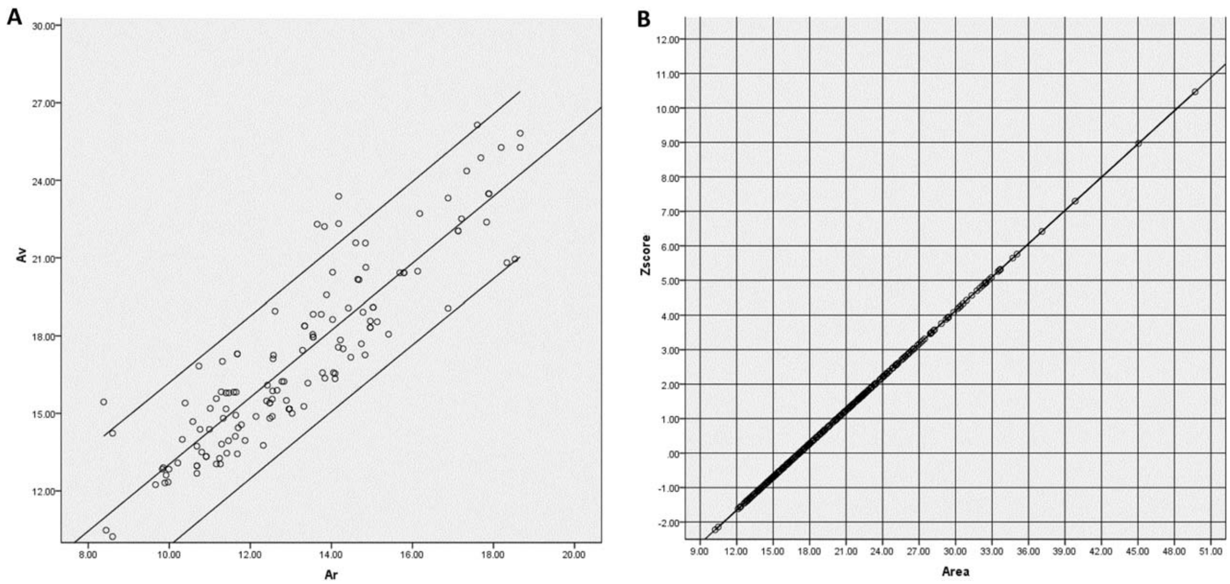


Figure 2. A, Scatterplot showing the 95% reference range and regression equation (Av = 0.096 + 1.294 × Ar) for the hiatal area on the Valsalva maneuver in 134 nulliparas. **B**, Diagram showing calculations of Z scores for the hiatal area on the Valsalva maneuver.



significant differences in the measurements and Z scores between POP stages ($F = 82.19$; $P < .001$; Table 1).

An ROC curve analysis was performed, and the curves of the measured values were perfectly consistent with the curves of Z-Av on the basis of substantial POP clinically and on US. The ROC curve analysis yielded areas under the curve of 0.69, 0.87, and 0.86 for the measurements and Z-Av values for POP-Q stage 2 or higher, substantial POP on translabial US, and symptomatic POP. The optimal Z-Av cutoff was 1, with sensitivity of 77% and specificity of 60% for POP-Q stage 2 or higher (Figure 4A), sensitivity of 84% and specificity of 75% for substantial POP on translabial US (Figure 4B), and sensitivity of 84% and specificity of 72% for symptomatic POP (Figure 4C). The levator hiatal area cutoff was 20 cm^2, with sensitivity of 81% and specificity of 58% for POP-Q stage 2 or higher, sensitivity of 86% and specificity of 70% for symptomatic POP, and sensitivity of 87% and specificity of 73% for substantial POP on translabial US.

Discussion

The levator hiatus is defined as the space enclosed by the levator ani muscle, which is the most important supporting tissue of the pelvic floor. To some extent, the levator hiatal dimensions reflect the biometric characteristics of the levator ani muscle.[1,2,15,16] Hiatal ballooning has been popularly expressed as percentages or absolute values of the changes in the hiatal area during the Valsalva maneuver.[8,17] The hiatal area is a repeatable diagnostic parameter. Its clinical application could improve our understanding of the pathophysiologic characteristics of POP as a form of hiatal hernia.[2,6]

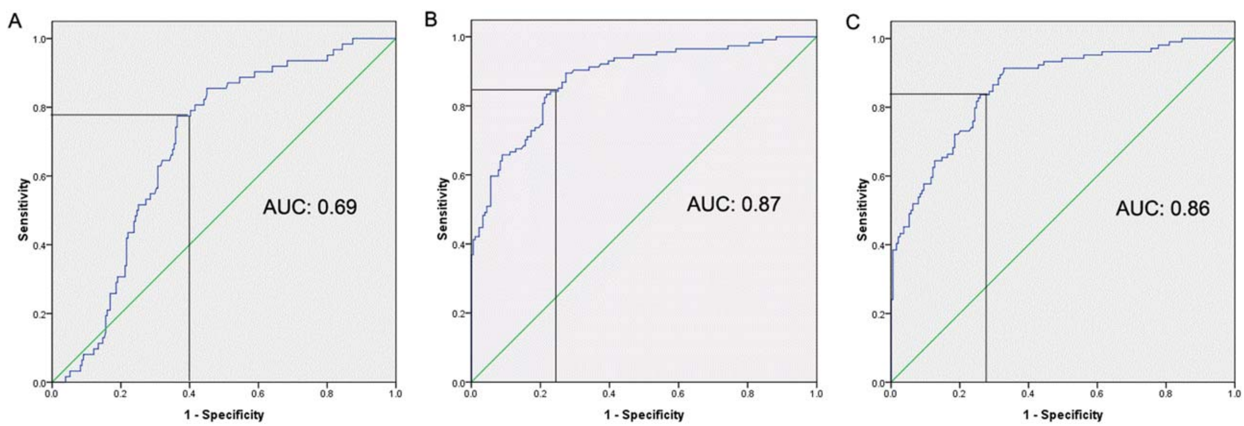
Our findings confirmed that the levator hiatal area on the Valsalva maneuver had a linear correlation with the levator hiatal area at rest ($r = 0.892$) in nulliparas. The maximal levator hiatal area on the Valsalva maneuver in nulliparas can be predicted on the basis of the hiatal area at rest by using regression equation and can be

Table 1. Analysis of Variance of the Levator Hiatal Area on the Valsalva Maneuver Against the ICS POP Stages (n = 293)

POP Stage	Area, cm^2	Z Score	F[a]	P[a]
0 (n = 134)	17.13 ± 3.50 (10.23–26.14)	0 ± 1.0 (−2.23–2.89)	82.19	<.001
1 (n = 46)	19.35 ± 3.42 (13.70–26.87)	0.7 ± 1.1 (−1.11–3.13)		
2 (n = 62)	22.93 ± 4.16 (13.98–33.49)	1.86 ± 1.34 (−1.02–5.25)		
3 (n = 43)	28.23 ± 4.79 (18.20–37.11)	3.56 ± 1.54 (0.34–6.42)		
4 (n = 8)	32.80 ± 10.85 (20.14–49.70)	5.03 ± 3.49 (0.96–10.47)		

Data are presented as mean ± SD (range).
[a]One-way analysis of variance.

Figure 4. Receiver operating characteristic curve analysis showing cutoffs for Z scores of the hiatal area on the Valsalva maneuver for the detection of ICS POP-Q stage 2 or higher (**A**), substantial POP on translabial US (**B**), and prolapse symptoms (**C**). AUC indicates area under the curve.



further transformed into Z scores, which exactly described the levator hiatal distensibility in this study. The diagrams were produced to simplify the calculation.

It seemed that the 95% reference range of the levator hiatal area on the Valsalva maneuver, which was derived from 134 nulliparas, provided a normal levator hiatal area reference scope. However, the ROC curve analysis illustrated that no diagnostic cutoff value can be used to distinguish the normal state and POP stage 1. Patients with POP stage 1 were always clinically asymptomatic.[18,19] Most of the levator hiatal areas were limited within the 95% reference range. If we consider women who had no POP symptoms to be "normal" and those with symptomatic POP to be "abnormal," the ROC curve analysis proposed a valuable outcome. A maximal Z-Av of 1.0 or levator hiatal area of 20 cm^2 resulted in high specificities and sensitivities. This finding suggested that the nulliparas who had a Z score above the normal scope (ie, ≥ 1) were likely to have an excessive levator hiatal distensibility. Further studies might determine whether this finding is a risk or protective factor of POP.

For women with no abnormalities, Z scores can be used to estimate whether the levator ani muscle distensibility or hiatal expansion is stronger or weaker than the general level under stretching conditions. For example, if the levator hiatal area of the nulliparas was measured to be 18 cm^2 at rest, which is greater than the average value, the individual would be predicted to have a maximal levator hiatal area on the Valsalva maneuver of 23.4 cm^2 under normal conditions, as calculated by the regression equation $Av = 0.096 + 1.294 \times Ar$. Meanwhile, a normal hiatus, with an area of 11 cm^2, possibly could expand to 14.3 cm^2. There is no doubt that both nulliparous hiatuses are normal. If the measured value is greater than the predicted value, it suggests that the distensibility is higher than the general level and vice versa. However, an Av cutoff of 20 cm^2 or greater may indicate that a value of 23.4 cm^2 is abnormal, whereas a value between 14.3 and 20 cm^2 is normal. Obviously, the Z score provided more information than the measured value regarding the evaluation of the distensibility of the levator hiatus.

For women with POP, the Z scores increased as the POP-Q stage increased; therefore, it seems plausible to define mild ballooning as a Z-Av of 1 to 3, moderate ballooning as a Z-Av of 3 to 5, marked ballooning as a Z-Av of 5 to 7, and severe ballooning as a Z-Av of 7 or greater. These definitions mean that mild ballooning indicated hiatal expansion that was in excess of 1 to 2 SDs over the mean level, and moderate and marked ballooning indicated that the hiatal expansion was an extra 2 to 4 SDs and 4 to 6 SDs beyond the normal level, respectively. The degrees of levator hiatal expansion were precisely and simply expressed. For example, an Av value of 25 cm^2 was equal to a Z score of 2.5, indicating that the hiatal expansion was in excess of 1.5 SDs over the mean (ie, mild ballooning).

We applied the Z score in a specific population using the same method. Our results confirmed the validity of this method from the clinical and research perspectives. However, there were, several limitations of this study that need to be acknowledged. First, it would be better to using a larger sample to calculate the Z score, which requires the population mean and population standard deviation. The standard deviation was estimated by using a random sample, since it is impossible to measure every member of a population. Second, the levator hiatal area measured by 3-dimensional US varied greatly in the population and had significant ethnic differences.[20,21] The applicability of the Z score would only be possible in specific populations whose normality curves are known, with the need to create specific population curves for clinical applicability. Our findings were based on a sample of patients that was limited to a population of Chinese women. In addition, because it is a new method, it is necessary to confirm its practicability and feasibility by further research.

In conclusion, the Z scores showed advantages in the evaluation of levator hiatal distensibility. The 95% reference range of the levator hiatal area on the Valsalva maneuver in Chinese nulliparas was expressed as Z scores between -2 and $+2$. A normal hiatal area on the Valsalva maneuver was defined as a Z-Av of less than 1.0. Mild ballooning was defined as a Z-Av of 1 to 3, moderate ballooning as a Z-Av of 3 to 5, marked ballooning as a Z-Av of 5 to 7, and severe ballooning as a Z-Av of 7 or greater. These findings also indicated that the nulliparous hiatus may have excessive distensibility, which is probably a potential risk or protective factor for POP. We intend to explore this point further, since little evidence has been shown in the literature.

References

1. Abdool Z, Dietz HP, Lindeque BG. Prolapse symptoms are associated with abnormal functional anatomy of the pelvic floor. *Int Urogynecol J* 2017; 28:1387–1391.

2. Khunda A, Shek KL, Dietz HP. Can ballooning of the levator hiatus be determined clinically? *Am J Obstet Gynecol* 2012; 206:246.e1–246.e4.
3. Dietz HP, Shek KL, Clarke B. Biometry of the pubovisceral muscle and levator hiatus by three-dimensional pelvic floor ultrasound. *Ultrasound Obstet Gynecol* 2005; 25:580–585.
4. Dietz HP, Shek C, De Leon J, Steensma AB. Ballooning of the levator hiatus. *Ultrasound Obstet Gynecol* 2008; 31:676–680.
5. Vergeldt TF, Notten KJ, Weemhoff M, et al. Levator hiatal area as a risk factor for cystocele recurrence after surgery: a prospective study. *BJOG* 2015; 122:1130–1137.
6. Dietz HP. Pelvic floor ultrasound: a review. *Clin Obstet Gynecol* 2017; 60:58–81.
7. Albrich SB, Welker K, Wolpert B, et al. How common is ballooning? Hiatal area on 3D transperineal ultrasound in urogynecological patients and its association with lower urinary tract symptoms. *Arch Gynecol Obstet* 2017; 295:103–109.
8. Shek Kl, Krause HG, Wong V, Goh J, Dietz HP. Is pelvic organ support different between young nulliparous African and Caucasian women? *Ultrasound Obstet Gynecol.* 2016; 47(6):774–778.
9. Royston P, Wright EM. How to construct "normal ranges" for fetal variables. *Ultrasound Obstet Gynecol* 1998; 11:30–38.
10. Min J, Wang VH, Xue H, Mi J, Wang Y. Maternal perception of child overweight status and its association with weight-related parenting practices, their children's health behaviours and weight change in China. *Public Health Nutr* 2017; 6:1–8.
11. Son MBF, Gauvreau K, Kim S, et al. Predicting coronary artery aneurysms in Kawasaki disease at a North American center: an assessment of baseline *z* scores. *J Am Heart Assoc* 2017; 6:pii:e005378.
12. Wen L, Zhang J, Zen S, Zhou Q. Using Z-scores to evaluate levator hiatal dimensions with four-dimensional translabial ultrasound. *J Obstet Gynaecol Res* 2017; 43:1840–1847.
13. Bump RC, Mattiasson A, Bø K, et al. The standardization of terminology of female pelvic organ prolapse and pelvic floor dysfunction. *Am J Obstet Gynecol* 1996; 175:10–17.
14. Dietz HP, Kamisan Atan I, Salita A. Association between ICS POP-Q coordinates and translabial ultrasound findings: implications for definition of "normal pelvic organ support." *Ultrasound Obstet Gynecol* 2016; 47:363–368.
15. Deegan EG, Stothers L, Kavanagh A, Macnab AJ. Quantification of pelvic floor muscle strength in female urinary incontinence: a systematic review and comparison of contemporary methodologies. *Neurourol Urodyn* 2018; 37:33–45.
16. Cheung RY, Shek KL, Chan SS, Chung TK, Dietz HP. Pelvic floor muscle biometry and pelvic organ mobility in East Asian and Caucasian nulliparae. *Ultrasound Obstet Gynecol* 2015; 45:599–604.
17. Andrew BP, Shek KL, Chantarasorn V, Dietz HP. Enlargement of the levator hiatus in female pelvic organ prolapse: cause or effect? *Aust NZ J Obstet Gynaecol* 2013; 53:74–78.
18. Gerges B, Kamisan Atan I, Shek KL, Dietz HP. How to determine "ballooning" of the levator hiatus on clinical examination: a retrospective observational study. *Int Urogynecol J* 2013; 24: 1933–1937.
19. Rostaminia G, White D, Hegde A, Quiroz LH, Davila GW, Shobeiri SA. Levator ani deficiency and pelvic organ prolapse severity. *Obstet Gynecol* 2013; 121:1017–1024.
20. Derpapas A, Ahmed A, Vijaya G, et al. Racial differences in female urethral morphology and levator hiatal dimensions: an ultrasound study. *Neurourol Urodyn* 2012; 31:502–507.
21. Abdool Z, Dietz HP, Lindeque BG. Interethnic variation in pelvic floor morphology in women with symptomatic pelvic organ prolapse [published online ahead of print June 17, 2017]. *Int Urogynecol J.* doi:10.1007/s00192-017-3391-7.